Breaking the Cycle?

Intergenerational Effects of an Anti-Poverty Program in Early Childhood *

Andrew Barr† & Chloe R. Gibbs‡

January 2018

Abstract

Despite substantial evidence that resources and outcomes are transmitted across generations, there has been limited inquiry into the extent to which anti-poverty programs actually disrupt the cycle of bad outcomes. We explore how the effects of the United States' largest early childhood program transfer across generations. We leverage the geographic rollout of this federally funded, means-tested preschool program to estimate the effect of early childhood exposure among mothers on their children's long-term outcomes. We find evidence of intergenerational transmission of effects in the form of increased educational attainment, reduced teen pregnancy, and reduced criminal engagement in the second generation.

Keywords: intergenerational, early childhood, Head Start, long-term

*We are grateful for financial support from the University of California, Davis Center for Poverty Research's Small Grants Competition, made possible in part by grant number AE00102 from the U.S. Department of Health and Human Services, Office of the Assistant Secretary for Planning and Analysis. We thank Maria Rosales-Rueda and Terri Sabol for helpful discussant comments and seminar participants at the University of Chicago Harris School, Stanford University Center for Education Policy Analysis, Institute for Research on Poverty Summer Research Workshop (Wisconsin), UC Davis, and W.E. Upjohn Institute for Employment Research for useful feedback. We also thank Martha Bailey, Marianne Bitler, Jason Lindo, Jens Ludwig, and Doug Miller for insightful conversations. Sreeraahul Kancherla and David Melnick provided excellent research assistance. Any remaining errors are our own. The views expressed in publications resulting from supported research do not necessarily reflect the official policies of the U.S. Department of Health and Human Services.

†Department of Economics, Texas A&M University, abarr@tamu.edu

‡Department of Economics, University of Notre Dame, chloe.gibbs@nd.edu.

1 Introduction

The effects of poverty are pernicious and persistent across generations. Those born to parents in the lowest quintile of the income distribution are twice as likely to end up there as children born to middle-income parents, and the intergenerational correlations in income, education levels, female headship, receipt of government assistance, and risky behavior are quite high.[1] Family, school, and neighborhood contexts collectively shape children's trajectories and generate correspondence between their parents' outcomes and their own. These linkages are particularly acute for minorities, potentially contributing to the early emergence and persistence of achievement gaps by race/ethnicity.[2]

Societal investments in education may disrupt the transmission of poverty across generations by increasing educational attainment and labor market attachment and decreasing engagement in risky behavior. Early childhood in particular is a critical developmental period and an opportunity for especially effective intervention. Indeed, multiple studies indicate that interventions in the preschool and early school years can have substantial effects on schooling attainment, labor market success, and other measures of health and well-being into adulthood.[3] And yet, we know almost nothing about whether these benefits carry over to the next generation. In other words, do these needs-targeted early childhood programs truly break the cycle of poverty?

We answer this question in the context of the Head Start program, providing the only evidence of which we are aware of the intergenerational effects of an early childhood intervention in the United States. The Head Start program, funded and administered through the U.S. Department of Health and Human Services, has been an integral part of the conversation about early childhood

[1] A variety of recent estimates suggest intergenerational correlations in income of 0.3 to 0.6 (Black and Devereux 2011, Chetty, Hendren, Kline, Saez and Turner 2014b, Mazumder 2005, Solon 1999), in education levels of 0.4 to 0.5 (Hertz et al. 2008), in female headship of 0.2 (Page 2004), and in welfare use of 0.3 (Page 2004). Similarly, Duncan and coauthors (2005) review and provide a variety of evidence indicating positive intergenerational correlations in early pregnancy, drug use, and other measures of delinquent or risky behavior.

[2] While there is some dispute about the magnitude of these gaps, there is consistent evidence that cognitive test-score gaps by race and ethnicity exist at formal school entry and remain throughout the schooling years (Fryer and Levitt 2004, Fryer and Levitt 2006, Murnane, Willett, Bub and McCartney 2006), and while race/ethnicity achievement gaps have narrowed in recent decades, gaps by socioeconomic status are pronounced (Reardon and Portilla 2016).

[3] Long-term evidence from the Abecedarian Project, Perry Preschool Project, Head Start, and the Project STAR class- size reduction intervention all suggest large positive effects on participants (Campbell et al. 2014, Chetty et al. 2011; Deming 2009, Dynarski et al. 2013, Garces et al. 2002, Heckman et al. 2013, Schweinhart et al. 2005).

intervention for the 50 years of its existence. Easily the largest early childhood education program in the United States, annual Head Start enrollment has grown from 400,000 during the early years of the program to nearly a million participants today. Across multiple datasets and different study designs, a large body of quasi-experimental evidence consistently indicates that participation in the Head Start program yields important long-term benefits, particularly for early cohorts of program participants (Carneiro and Ginja 2014, Deming 2009, Garces, Thomas and Currie 2002, Ludwig and Miller 2007).[4]

To explore intergenerational spillovers, we capitalize on differential exposure to Head Start induced by variation in the early rollout of the program. We generate Head Start availability measures using data we extracted and compiled from the National Archives and Records Administration (NARA).[5] We link these measures with the National Longitudinal Survey of Youth–1979 Cohort (NLSY79) and the NLSY79 Children and Young Adults Survey (CNLSY) in order to compare the children of mothers who differ in terms of Head Start availability. We are interested in effects on the second generation's long-term outcomes including educational attainment, teen pregnancy, and criminal engagement.

We find a significant impact of Head Start availability, between 0.25 and 0.45 standard deviations, on a summary index of long-term outcomes for the second generation. These estimates are robust to the inclusion of a variety of flexible controls for within-county and within-state variation across birth cohorts, including county by birth year trends, as well as direct controls for the time-varying availability of other War on Poverty programs within counties. The legitimacy of the geographic rollout strategy is bolstered by estimates that demonstrate no relationship between Head Start availability and the outcomes of children unlikely to have been eligible for the program.

Our findings indicate that societal investments in early childhood education can disrupt the intergenerational transmission of the effects of poverty. Indeed, when comparing children of mothers more or less likely to have grown up in poverty, our estimates suggest that Head Start closes most

[4] Estimates of the effect of Head Start on more recent cohorts of participants is less clear. Results of the National Head Start Impact Study, the first large-scale, randomized controlled study of the program showed initial impacts on cognitive and non-cognitive skills for Head Start participants, but these effects faded almost entirely by the first and third grades (Puma et al. 2005, Puma et al. 2010, Puma et al. 2012). However, quasi-experimental evidence indicates meaningful effects on long-term outcomes despite the fade out of test-score impacts (Deming 2009).

[5] See Appendix C for details on construction of the NARA data.

of the gap in a summary index of long-term outcomes for the second generation.

2 A Path Out of Poverty

There is substantial evidence documenting the path dependency of socioeconomic status. At each stage from school readiness through adulthood, children from the highest family income quintile are dramatically more likely to obtain benchmarks of lifetime economic success than those from the lowest quintile (Sawhill, Winship and Grannis 2012). While children growing up in disadvantage fall behind at each stage, successful navigation of each stage from early childhood to adolescence translates into increased likelihood of attaining a middle class existence (Sawhill et al. 2012). These effects carry over to the next generation, resulting in high intergenerational correlations in income, educational attainment, and risky behavior.

Evidence on the lack of intergenerational mobility in the United States, coupled with growing income inequality, has led to considerable interest in understanding why the resources, behaviors, and outcomes of parents are so strongly related to those of their children, and furthermore, whether interventions that improve these behaviors and outcomes might carry over to the affected individuals' children (Auten, Gee and Turner 2013, Chetty et al. 2014b, Corak 2013, Lee and Solon 2009). The existing evidence on the collective importance of childhood contexts—families, schools, and neighborhoods—in determining longer-term outcomes suggests that social interventions or programs that intervene in these contexts may be influential (Chetty, Hendren, Kline and Saez 2014a).

Despite the clear importance of this question, the data requirements necessary to answer it convincingly have resulted in limited applications. While we are unaware of any study to focus on the long-term intergenerational effects of an anti-poverty program, several have estimated the intergenerational effects of increases in educational attainment in adolescence and beyond. Increases in college access or attainment have resulted in improved birth outcomes and reduced grade retention in the next generation (Currie and Moretti 2003, Maurin and McNally 2008, Page 2009). The evidence at the middle and high school levels is more mixed, with positive intergenerational effects of additional schooling generated by compulsory schooling changes in the U.S. and Great Britain in the 1960s and 70s and no effect in Norway (Black, Devereux and Salvanes 2005, Oreopoulos, Page and Stevens 2006, Chevalier 2007). Taken together, the evidence for the intergenerational effects of education is promising. However, we know little about the intergenerational effects of early childhood programs, despite the large estimated effects of these types of programs on adult outcomes.[6]

Indeed, there is a substantial body of empirical evidence demonstrating that early childhood programs can generate improvements in participants' life chances over the long-term, and, furthermore, evidence that early skills are important predictors of subsequent academic attainment and labor market success (Chetty et al. 2011, Duncan et al. 2007, Dynarski, Hyman and Schanzenbach 2013). Specifically, evidence from the Abecedarian Project, Perry Preschool Project, Head Start, and the Project STAR class-size reduction intervention collectively suggests that interventions in the preschool and early school years can have substantial effects on schooling attainment, labor market success, and other measures of well-being into adulthood (Chetty et al. 2011, Deming 2009, Schweinhart et al. 2005). Recent evidence documents improvements in life chances that include better health, reductions in behavior problems, and higher rates of college-going (Campbell et al. 2014, Carneiro and Ginja 2014, Dynarski et al. 2013).

We contribute to this important conversation by providing some of the first evidence—and the first evidence in a U.S. context, to our knowledge—on whether the effects of early childhood programs transfer across generations. The answer to this question has important implications for policies aimed at reducing poverty or socioeconomic gaps in educational attainment, risky behaviors, and relatedly, labor market success. Most importantly, if these types of policies have large spillover effects on the next generation it suggests a concerted effort for a single generation of impoverished youth might break the cycle of poverty and reduce the need to provide similar services to future generations.

[6] Rossin-Slater and Wust (2017) explore the intergenerational impact of a Danish preschool program and a nurse home visiting program in infancy. They find educational attainment effects in the first generation that persist, and are similarly-sized, in the second generation. We return to discussion of this study later in the paper.

2.1 The Evidence on Head Start

Recent policy discussions of large-scale, publicly provided preschool interventions often rely on the Head Start literature as most relevant and informative for designing and scaling up programs. The Head Start program was an early piece of President Lyndon B. Johnson's War on Poverty, commencing as a summer program in 1965, serving 560,000 children (Vinovskis 2005). It was then quickly expanded to a year-round program. While Head Start today is characterized as an early childhood education program, it was designed as an anti-poverty program with significant health and community development components. The initial emphasis was on a variety of "preschool"-related services and supports, including nutrition, vaccinations and health care, dental services, and social development (Vinovskis 2005).

In its present-day form, the Head Start program is not considered as high-quality as the much-touted Perry Preschool and Abecedarian Projects—in part because of its somewhat lower per child funding levels—but Head Start began as a comprehensive attempt to aid poor children. In the early years of the program, the mission of Head Start was characterized as "providing the children of the poor with an equal opportunity to develop their full potential" (Office of Child Development 1970). To this end, centers provided medical and dental care, nutritional services, parent involvement activities, employment and training to the disadvantaged, linkages to social services, and mobilization of community resources in addition to providing programming to foster children's social-emotional and cognitive development.[7]

Head Start served a decidedly disadvantaged population in the early years of the program. The median family income was less than half that of all families in the U.S. and approximately 50 percent of early full-year program participants were black children (Office of Child Development 1968). In the early years of the program, between nine and 17 percent of families reported having no running water inside the home. Only five percent of mothers reported some postsecondary schooling or more, with approximately 25 percent indicating that they graduated from high school, and 65 to 70 percent of mothers with less than a high school education. Approximately 25 percent lived in female-headed households and between 65 and 70 percent of participating children's mothers were

[7] Depending on the year, 40 to 50 percent of families indicated that their child's Head Start medical examination uncovered a problem, and of those found to have an issue, 65 to 70 percent received treatment.

unemployed (Office of Child Development 1968).

While there is some debate about the pattern of short-run Head Start effects, prior quasi-experimental studies suggest Head Start has had important long-term effects for cohorts of children who participated from the late 1960s through the 1980s.[8] Leveraging sibling comparisons and discontinuities in grant-writing assistance and program eligibility, studies have documented increased educational attainment, better health, higher earnings, and less involvement in risky behaviors (Carneiro and Ginja 2014, Deming 2009, Garces et al. 2002, Ludwig and Miller 2007), even in the presence of short-term test-score fadeout (Deming 2009). A recent follow-up to Deming's study looks at even longer-term outcomes and finds persistence of effects later into adulthood, including impacts on participants' later-life parenting practices (Bauer and Schanzenbach 2016).

To build on the existing evidence of Head Start effectiveness and extend our understanding of the intergenerational effects of anti-poverty efforts, we capitalize on plausibly random exposure to the early rollout of the Head Start program over geography and birth cohorts (Figure 1). The Head Start program was rolled out quickly as a featured, and politically popular, component of the War on Poverty (Zigler and Valentine 1979). Announced in President Lyndon B. Johnsons 1965 State of the Union address in January, the program was operational that summer. Grant funds were distributed directly to local grantees as a means to circumvent governors, state legislatures, and agencies that may have prevented the funds from reaching disadvantaged black children (Gibbs, Ludwig and Miller 2011, Vinovskis 2005). In the early years of the program, approximately 40 percent of counties in the U.S. received Head Start funding. As a result of the local distribution of funding, programs became available in different counties at different times. We leverage this variation to identify the intergenerational effects of Head Start availability.

Our strategy is similar to the novel approach employed in looking at the impact of other War on Poverty and poverty reduction programs, including the Food Stamp Program (Almond, Hoynes

[8] While the Head Start Impact Study (HSIS) found initial positive effects on cognitive skill for participants in the mid-2000s, there were no persistent effects at first and third grade follow-ups (Puma et al. 2005, Puma et al. 2010, Puma et al. 2012). Re-analyses of the HSIS data suggest a more nuanced picture (Montialoux 2016). These analyses revealed that there is considerable variation in impact by center (Walters 2015), that effects are most pronounced among children who would otherwise be in parental or relative care (Kline and Walters 2016), and that Hispanic children and children with low skills at program entry experience the greatest benefit (Bitler, Hoynes and Domina 2014).

and Schanzenbach 2011, Hoynes, Schanzenbach and Almond 2016), the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC) (Hoynes, Page and Stevens 2011), and Community Health Centers (Bailey and Goodman-Bacon 2015). Two concurrent papers use the early introduction of Head Start over geography and time to explore impact on first-generation outcomes; preliminary evidence in both cases suggests sizable, long-term impact for participants (Johnson and Jackson 2017, Thompson 2017), particularly when coupled with subsequent schooling investment (Johnson and Jackson 2017).[9]

3 Data

To explore Head Start's intergenerational effects we rely on rich, longitudinal survey data that connect mothers and their children. The NLSY79 is a nationally representative sample of adolescents who were 14 to 22 years old when they were first surveyed in 1979. The survey follows 12,686 young men and women, with annual interviews through 1994 and biennial interviews continuing since then. In addition to rich data on labor market participation and transitions, the data provide extensive information on education and training, health, mobility, and family formation. The data facilitate analysis on a representative sample of young American men and women living in the United States in 1979. The timing is particularly advantageous for the purposes of this study as individuals born during the early 1960s are differentially exposed to Head Start via its introduction and rollout. Our analytic sample is restricted to NLSY79 respondents (mothers) born between 1960 and 1964.[10]

Beginning in 1986, a separate, related survey of all children born to NLSY79 female respondents has been collected, the CNLSY. In addition to all the mother's information from the NLSY79,

[9]Ludwig and Miller (2007) also rely on county-level Head Start exposure, leveraging a county poverty rate discontinuity in Head Start grant-writing assistance, to estimate program impact on first generation outcomes. As they note, this approach is challenging when using survey data as many counties are not represented at all and there are limited individual observations around the discontinuity.

[10]This is due to measurement and missingness issues with grandmother (mothers of NLSY79 respondents) education levels for pre-1960 cohorts. While maternal education levels for 1960 to 1964 cohorts correspond very closely with other data sources (Current Population Survey) both in levels and trends, there is a significant positive jump in education levels for those born prior to 1960, opposite the trend observed in the CPS. These issues are problematic as we use grandmother education levels, a proxy for Head Start eligibility, to restrict our sample to mothers likely to be eligible for Head Start.

the child survey includes direct information for each child collected from either the mother or child depending on age. The survey gathers data on children's schooling and training, labor market experiences, health, and engagement in risky behaviors. The CNLSY allows us to explore intergenerational effects of the mother's Head Start exposure.

In addition to the NLSY surveys, we use county-by-year data from the Community Action Programs (CAP) and Federal Outlays System (FOS) files available from the National Archives and Records Administration for Head Start availability in fiscal years 1966–1968 (see Data Appendix for details). We aggregate data on Head Start grant recipients to the county level, and code a birth cohort and county pair as "exposed" to Head Start if it received per four-year old Head Start funding above the tenth percentile ($22 per four-year old in the county).[11] We do not otherwise leverage data on appropriated dollar amount because of concerns about the accuracy of the recorded funding amounts in the early years of the Head Start program as well as the endogeneity of funding levels.

We focus on the second generation's long-term outcomes both because these outcomes are most important in assessing whether the intergenerational transmission of program effects disrupts the cycle of poverty and because there are myriad ways in which a mother's Head Start access may affect her children's outcomes. These pathways are likely cumulative across childhood. From changes in parenting practices and greater likelihood of enrolling one's child in early childhood programming to heightened expectations and spillovers from a mother's own increased human capital and income, we would expect the channels of impact on the second generation to accumulate over the childhood years. There are two positive outcomes: high school completion (including GED receipt) and college going, or attending college for any period of time. Because of findings in prior literature, we also consider two negative outcomes with important implications for children and teens' life changes: teen pregnancy and interaction with the criminal justice system (as measured by any arrests, convictions, or probations). These outcomes are important in capturing the second generation's private returns, but also have implications for measuring the broader societal benefits of the program. To address multiple inference concerns and reduce measurement error, we follow

[11] In other words, we classify county-birth cohorts with very low funding levels as unexposed. Our results are qualitatively similar using $0 as the cutoff.

the prior literature in constructing a summary index of our outcome measures (Kling, Liebman, and Katz 2007; Deming 2009). We normalize each outcome to have a mean zero and standard deviation one, adjust outcome signs so that a more positive outcome is better (i.e., we flip the sign on teen parenthood and crime), and take the simple average across these outcomes. We then standardize this summary index.

The top panel of Table 1 contains summary statistics for these outcomes. The first column provides these measures for the full sample. Columns (2) and (3) contain similar measures for the samples underlying our geographic rollout strategy. As participation in Head Start is largely restricted to children in poor households, we use a proxy for family resources (maternal education) to restrict our sample to individuals who are likely to have been affected.[12] Consistent with this approach, levels of maternal educational attainment among early Head Start participants were very low. We focus our analyses on the children of NLSY79 mothers whose mothers (i.e., the grandmothers of our population of interest) did not finish high school (column (2)).[13] We refer to this as our high impact sample because close to 70 percent of participants' mothers had less than a high school degree, implying participation rates of close to 60 percent in counties with Head Start availability.[14] As seen from the statistics, children in our high impact sample are negatively selected relative to the population. They have higher rates of teen parenthood (22% vs. 19%) and interaction with the criminal justice system (31% vs. 29%), and lower rates of high school completion (78% vs. 82%) and college going (51% vs. 58%). In column (3), we restrict the sample to NLSY79 mothers whose mother completed at most a high school degree. We call this our low impact sample as we estimate participation rates of at most 30 percent in counties with Head Start availability. As might be expected, summary statistics for children in this group have somewhat better outcomes.

We present analogous statistics for maternal outcomes (i.e., NLSY79 respondents) in the bot-

[12]The NLSY79 lacks measures of household income prior to the age of initial sample inclusion (in adolescence), so we focus our analyses on families that are likely to have been poor based on the education level of the NLSY79 respondent's mother, which is unaffected by Head Start availability for her child.

[13]Roughly 65 to 70 percent of the mothers of early participants reported less than a high school education, approximately 25 percent indicated that they graduated from high school, and only five percent of mothers reported some postsecondary schooling or more.

[14]See the note to Table A1 for details of these calculations.

tom panel of Table 1. While rates of teen parenthood are substantially higher and rates of college going are substantially lower, the pattern of statistics across samples is similar.[15]

4 Estimation and Empirical Results

To circumvent issues associated with individual selection into Head Start, we rely on variation within counties over time in the availability of the Head Start program. This variation is generated by the rollout of the Head Start program during the mid-1960s. We explore how a mothers exposure to the Head Start program affects the adult outcomes of her children.

We leverage the rollout of the Head Start program during the mid-1960s to identify the effect of program availability within a mother's county of birth on the adult outcomes of her children. As between 65 and 70 percent of the mothers of early Head Start participants did not complete high school (Barnow and Cain 1977, Office of Child Development 1968), we focus our analyses on this "high impact" population. Our basic specification is:

$$y_{ict} = \beta_0 + \beta_1 X_i + \beta_2 HSavail_{ct} + \gamma_c + \lambda_t + \varepsilon_{ict} \tag{1}$$

where y_{ict} is an adult outcome for a child; X_i includes controls for the child's sex and age as well as the mother's birth order and race; and γ_c and λ_t are county of birth and birth year fixed effects. $HSavail_{ct}$ indicates whether Head Start was available for a mother in a particular birth cohort t and birth county c. $HSavail_{ct}$ is set to one for a mother when there is a non-trivial level of Head Start funding in that mother's birth county four or five years after her year of birth.[16] Standard errors are clustered at the county of birth level.

Table A1 illustrates that our measure of Head Start availability predicts both self-reported Head Start participation and state-level participation rates derived from administrative Head Start enrollment data. The top panel presents estimates of the effect of Head Start availability on self-reported participation. When a program is available in a county four or five years after a

[15] Similar measures of interaction with the criminal justice system are unavailable for this sample.

[16] Explicitly, $HSavail_{ct}$ is set to one when the level of Head Start funding within a county exceeds the 10th percentile of observed funding per four year old (roughly \$22). The results are qualitatively similar using a cutoff of \$0.

mother's year of birth, the mother in our high impact sample is 10 percentage points more likely *to report* having participated in Head Start as a child. The Head Start participation question is asked retrospectively in 1994, when sample members were 30 to 34 years old. In other words, the participation measure reflects individuals recollection of whether they participated in a program nearly 30 years earlier, when they were 3 to 5. A variety of evidence from the psychology literature indicates that retrospective reports of early childhood are extremely unreliable (see Appendix B for further discussion). Given the self-reported and retrospective nature of the Head Start participation variable, we expect there is considerable misreporting.

As has been established in the literature, measurement error in a binary variable is necessarily non-classical and will thus result in a downwardly biased estimate of the relationship between Head Start availability and participation (Kane, Rouse and Staiger 1999, Aigner 1973). Under reasonable assumptions on the extent of misreporting, the true relationship between Head Start availability and Head Start participation may be four to eight times the observed estimate in our high impact sample, or 40 to 80 percentage points (see Appendix B for the details of these calculations). Given these concerns, we do not interpret these estimates as a first stage. Nevertheless, the positive relationship between program availability and self-reported participation supports our research design. The middle panel contains estimates using state-level variation in participation rates in 1966.[17] The estimates suggest close to a 30 percentage point increase in the likelihood of participation for individuals in our high impact sample. The table also contains implied participation rates, assuming that all participation occurred in counties with Head Start availability. These estimates provide a reasonable upper bound for the effect of Head Start availability on participation in our high impact sample of nearly 60 percentage points, similar to the estimated relationship between Head Start availability and participation after correcting for measurement error.

Our baseline results are contained in Table 2. The first row contains estimates from our high impact sample, restricted to grandchildren with grandmothers with less than a high school degree. We observe a large positive effect (0.47 sd) of a mother's Head Start availability on our index of adult outcomes for her children that is statistically significant at the 1 percent level. This effect

[17] Disaggregated state-level enrollment numbers are not available in other years during this time period. The note to Table A1 contains the details of these calculations.

is driven by significant reductions in teen parenthood (8 percentage points) and criminal behavior (15 percentage points) and increases in high school graduation (13 percentage points) and college enrollment (17 percentage points). The second row presents estimates from our low impact sample. As expected given their lower levels of Head Start participation and disadvantage, the effects are smaller for individuals in this group: a 0.22 sd increase on our index of adult outcomes.

While our baseline inference relies on standard errors clustered at the county of birth level, we have also explored the robustness of our p-values to the more conservative approach of randomization inference.[18] Under this procedure (essentially a large set of placebo assignments), we randomly reassign the pattern of Head Start introduction timing to counties of birth and estimate our basic specification. We do this 1,000 times. The distribution of these estimates is contained in Figures A1 and A2. As displayed in the figures, the estimates we observe are quite unlikely under a random assignment of the availability of Head Start to a county.[19] P-values presented are the two-tailed statistics calculated as the share of coefficient estimates obtained under random assignment of Head Start timing that are larger in absolute magnitude than the estimate produced using the true timing assignment. Our randomization inference p-values are similar to those obtained using a more standard approach.

4.1 Threats to Internal Validity

To interpret these estimates as the causal effect of Head Start availability, it must be the case that the availability of a Head Start program is, conditional on county and year of birth fixed effects, unrelated to other factors that would affect the outcomes of children born to women who did and did not have the program available. For example, one concern would be that the type of woman who became a mother or the type of woman included in the sample (due to non-response or our sample restrictions) was affected by the availability of a Head Start program in early childhood. To check for this we examine how Head Start availability predicts maternal background characteristics that are unlikely to have been affected by Head Start directly (columns (1)-(5) of

[18]See Abadie, Diamond, and Hainmueller (2010) for a discussion of this procedure and examples of its implementation.
[19]The random reassignment is at the level of county of birth.

Table A2). We do this exercise separately for the full sample and the two restricted samples we use for our rollout analyses. There is little relationship between maternal characteristics (race, maternal birth order, 1978 household poverty status) and Head Start availability. Similarly, there is no evidence that the education levels of the grandmother, which we use to focus our sample, are affected by Head Start availability. In column (6) and (7), we present analogous estimates focused on second generation characteristics unlikely to be affected by Head Start: the age and gender of the child. While there is no relationship with child age, Head Start availability is correlated with child gender in our high impact sample. Given the general balance of observables across samples and characteristics we are not particularly concerned by this difference, but as a further check we explore the relationship between Head Start availability and a predicted index based on all of the characteristics in the table.[20] We would be concerned if there were a positive and significant coefficient. Instead, we see insignificant negative point estimates, suggesting that, if anything, the children of those who have Head Start available are likely to have worse outcomes based on exogenous observable characteristics. Finally, we explore how our treatment estimates are affected by the inclusion of covariates. As expected from the balance of observable characteristics, our estimates are robust to the exclusion of covariates, supporting the argument that the availability of Head Start is conditionally exogenous.[21]

A second concern is that of endogenous program adoption or pre-existing positive trends in outcomes in counties that adopted a Head Start program. The presence of meaningful pre-trends might reflect general improvements in early childhood conditions or the existence of other programs that were correlated with Head Start availability. While the relatively tight window of analysis limits concerns related to pre-existing trends, we also probe the robustness of our estimates to the inclusion of differential trends by birth county (Table 3). Allowing differential birth cohort trends interacted with baseline (1960) county characteristics, or even birth county-specific trends, does

[20]To implement this approach, we first regress the index on all of the maternal and child characteristics in Table A2. Nearly all of the covariates are significant predictors at the 1% level and the p-value from an F-test in this regression is 0.000, indicating that these covariates are predictive of our index of outcomes. We then regress this predicted index on Head Start availability and county and year of birth fixed effects.

[21]Concerns about differential selection of families into the sample are further limited by the inclusion of family fixed effects into the specification. While this approach has a number of limitations, the resulting point estimates (in column (8)) are statistically indistinguishable from the other point estimates in Table 3.

little to change our point estimates. The estimates are similarly robust to the inclusion of more specific county-cohort controls for spending on War on Poverty programs and state by birth cohort fixed effects, which flexibly control for changes over time within states that could affect maternal outcomes.[22]

Figure 2 addresses these concerns graphically, demonstrating the relationship between Head Start availability within a county and the index of adult outcomes for the second generation. The x-axis presents the number of years between a mothers year of birth and the first year of Head Start availability in a county. Those individuals with a non-negative value are considered treated in that Head Start was available in their county of birth. As observed in the figure, the estimates are flat and close to zero prior to the availability of Head Start and then positive after a program becomes available, consistent with our estimates representing a causal effect of Head Start availability.

We further address endogeneity concerns related to the availability of a Head Start program using a placebo exercise. In Table 4, we explore the effect of Head Start availability on the children of a group of individuals who are largely ineligible for the program. Specifically, we run our basic specification on the children of mothers whose mothers obtained at least a high school degree. Only a small fraction of women in this group were eligible for or participated in Head Start.[23] If something other than Head Start availability is driving our main results, we might expect to see similar affects show up for the children of women in this group. Table 4 illustrates that the point estimates for this group are small, frequently opposite-signed, and non-significantly different from zero across all outcomes.[24]

4.2 Effect Size and Heterogeneity

As with much research estimating effects of early childhood programs, the effects are substantial (Deming 2009, Garces et al. 2002, Heckman, Pinto and Savelyev 2013, Johnson and Jackson 2017).

[22]The War on Poverty program variables control for spending on Medicaid, Community Action Program (CAP) administrative grants, cash assistance, CAP health programs, and Community Health Centers.

[23]We estimate that participation rates in this group were at most 1/5 of those in our high impact sample. Appendix Table A3 further illustrates that our measure of Head Start availability has a small and non-significant relationship with self-reported Head Start participation in this subsample.

[24]We similarly find non-significant point estimates in the sample with grandmothers who obtained more than a high school degree, but the sample sizes are too small for this exercise to provide much information. For example, the estimated effect of Head Start availability on the index of outcomes is -0.204 (se 0.388) for this group.

Furthermore, where comparable, the effects on the second generation are at least as large as effects on the first generation (Table 5). We cannot reject the equivalence of effect sizes across most measures. While the level effects are generally larger in the second generation, the percentage effects are similar for most education measures.[25] While there are few benchmarks for comparison, this high intergenerational correlation in effects is consistent with some recent findings (Rossin-Slater and Wust 2016). Rossin-Slater and Wust (2016) show that the children of women with access to a government-approved preschool are 1.2 percentage points more likely to have more than a compulsory education by age 25. This effect is similar in magnitude to the effect in the first generation (1.3 percentage points).

Notably, research on the effects of childhood health care access, food assistance, and education programs in a similar time period has documented large long-term impacts on the first generation (Almond et al. 2011, Goodman-Bacon 2016, Heckman et al. 2013, Johnson and Jackson 2017). To put our results in the context of recent literature with similar outcome measures, our implied TOT effects on high school graduation (21-25 percentage points) for the second generation are similar to those estimated for participants in the Perry Preschool program (20 percentage points), those estimated previously for Head Start (20 percentage points for white participants), and about a third of the size of those estimated for Food Stamp usage in families with young children (74 percentage points).[26] While it is difficult to construct comparable measures for criminal behavior, our implied TOT effects on criminal behavior (26 and 18 percentage points) are larger than the effects on somewhat similar measures reported in evaluations of the Perry Preschool program (12 percentage points on arrest by age 27, 19 percentage points on five or more arrests by age 27) and Head Start (12 percentage points for blacks). However, given power limitations in all of these studies, the confidence intervals are overlapping. Furthermore, Perry Preschool enrolled a very particular type of student, extremely disadvantaged, black children in Ypsilanti, Michigan. While

[25]We do not include the estimate on highest grade completed in the second generation in our main tables as many of these individuals have not finished their education; however, the point estimates are 0.670 (se 0.225) for the high-impact sample and 0.284 (se 0.175) for the low-impact sample, statistically indistinguishable from the effect on the mothers.

[26]Our estimates are also relatively similar to Johnson and Jackson (2017) who find that Head Start availability leads to an eight percentage point increase in the likelihood of finishing high school for likely participants, quite similar to our first-generation estimates (Table 5).

statistically indistinguishable from our full sample estimates, if we use our estimates for the children of black mothers, our implied TOT estimates are roughly half the size of the estimates from the Perry Preschool evaluations across comparable measures (Table 6).[27] Our effects show up across subgroups (Table 6), with somewhat larger effects on crime and somewhat smaller effects on teen parenthood for male children, perhaps as a result of the higher rate of crime and lower rate of reported teen parenthood for this group.

Of course it may not be reasonable to convert our estimates to TOT effects as there may be important spillover effects of program availability; indeed, it is not difficult to imagine that improving the trajectories of a significant share of a group results in improvements for the group as a whole that are substantially larger than what we might expect to see if an individual was treated in isolation (as in the Perry Preschool experiment where only 58 children were offered a place in the program). To the extent that subsequent schooling quality or productivity increased because a share of each school-entry cohort was more prepared, these sorts of spillovers could occur. Given this, we focus on the estimated effect of Head Start availability.

5 Discussion and Conclusion

Research and policy discussion frequently focus on how to level the playing field for those born into poverty. We focus instead on whether such interventions truly break the cycle of bad outcomes. While there is increasing interest among researchers and policymakers in understanding the intergenerational spillovers of particular policies and interventions, there exist very few contexts in which these questions can be tested empirically; this is particularly true for early childhood interventions, an area of increasing focus and investment. The federal Head Start program provides a context in which data availability and the time horizon since first implementation facilitate such an empirical exploration, allowing us to contribute the first U.S. evidence on the intergenerational transmission of early childhood intervention effects.

We find consistent evidence that the positive effects of Head Start during its earliest years transferred across generations in the form of improved long-term outcomes for the second genera-

[27]This is due to the higher participation rates for black individuals as well as the somewhat smaller point estimates.

tion. The pattern of results suggests decreases in teen parenthood and criminal engagement and increases in educational attainment across empirical approaches. The effects are large in magnitude, but broadly consistent with the positive first-generation effect sizes found in evaluations of similar early childhood programs that provided an array of services to disadvantaged youth.[28] Furthermore, because of the large scale of Head Start, the program likely provided benefits beyond the direct effect on participants.

Indeed, the availability of Head Start, at least during the early years of the program, appears to have been quite successful at breaking the cycle of poor outcomes for disadvantaged families. Head Start access closes most of the gap in outcomes between individuals with more and less advantaged grandmothers. These results imply that cost-benefit analyses of Head Start and similar early childhood interventions underestimate the benefits of such programs by ignoring the transmission of positive effects across generations. This finding has important policy implications for optimal investment in these types of programs. Each disadvantaged child society helps now will lead to fewer who require assistance in the future.

[28] Recall that during the early years of the program Head Start provided medical and dental care in addition to engaging parents of the participants and providing linkages to other social services to an extremely disadvantaged population of children and families.

References

Aigner, Dennis J, "Regression with a binary independent variable subject to errors of observation," *Journal of Econometrics*, 1973, *1* (1), 49–59.

Almond, Douglas, Hilary W. Hoynes, and Diane Whitmore Schanzenbach, "Inside the War on Poverty: The impact of food stamps on birth outcomes," *Review of Economics and Statistics*, 2011, *93(2)*, 387–403.

Auten, Gerald, Geoffrey Gee, and Nicholas Turner, "Income inequality, mobility, and turnover at the top in the US, 1987-2010," *American Economic Review*, 2013, *103(3)*, 168–172.

Bailey, Martha J. and Andrew Goodman-Bacon, "The War on Poverty's experiment in public medicine: Community health centers and the mortality of older Americans," *American Economic Review*, 2015, *105(3)*, 1067–1104.

—— **and Nicolas J. Duquette**, "How Johnson fought the War on Poverty: The economics and politics of funding at the Office of Economic Opportunity," *Journal of Economic History*, 2014, *74(2)*, 351–388.

Barnow, Burt S. and Glen G. Cain, "A reanalysis of the effect of Head Start on cognitive development: Methodology and empirical findings," *Journal of Human Resources*, 1977, *12(2)*, 177–197.

Bauer, Lauren and Diane Whitmore Schanzenbach, "The long-term impact of the Head Start program," Economic Analysis, Brookings Institution, The Hamilton Project, Washington, DC 2016.

Bitler, Marianne P., Hilary W. Hoynes, and Thurston Domina, "Experimental evidence on the distributional effects of Head Start," Working Paper 20434, National Bureau of Economic Research 2014.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes, "Why the apple doesn't fall far: Understanding intergenerational transmission of human capital," *American Economic Review*, 2005, *95(1)*, 437–449.

Black, Sandy and Paul Devereux, "Recent developments in intergenerational mobility," *Handbook of Labor Economics*, 2011, *4*, 1487–1541.

Campbell, Frances, Gabriella Conti, James J. Heckman, Seong Hyeok Moon, Rodrigo Pinto, Elizabeth Pungello, and Yi Pan, "Early childhood investments substantially boost adult health," *Science*, 2014, *343*, 1478–1485.

Carneiro, Pedro and Rita Ginja, "Long-Term impacts of compensatory preschool on health and behavior: Evidence from Head Start," *American Economic Journal: Economic Policy*, 2014, *6(4)*, 135–173.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan, "How does your kindergarten classroom affect your earnings? Evidence from Project STAR," *Quarterly Journal of Economics*, 2011, *126(4)*, 1593–1660.

___ , Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the land of opportunity? The geography of intergenerational mobility in the United States," Working Paper 19843, National Bureau of Economic Research 2014.

___ , ___ , ___ , ___ , and Nicholas Turner, "Is the United States still a land of opportunity? Recent trends in intergenerational mobility," Working Paper 19844, National Bureau of Economic Research 2014.

Chevalier, Arnaud, "Parental education and child's education: A natural experiment," IZA Discussion Paper 1153, Institute for the Study of Labor 2007.

Corak, Miles, "Income inequality, equality of opportunity, and intergenerational mobility," *Journal of Economic Perspectives*, 2013, *27(3)*, 79–102.

Currie, Janet and Duncan Thomas, "Does Head Start make a difference?," *American Economic Review*, 1995, *85(3)*, 341–364.

—— **and Enrico Moretti**, "Mother's education and the intergenerational transmission of human capital: Evidence from college openings," *Quarterly Journal of Economics*, 2003, *118(4)*, 1495–1532.

Deming, David, "Early childhood intervention and life-cycle skill development: Evidence from Head Start," *American Economic Journal: Applied Economics*, 2009, *1(3)*, 111–134.

Duncan, Greg, Ariel Kalil, Susan E Mayer, Robin Tepper, and Monique R Payne, "The apple does not fall far from the tree," *Unequal chances: Family background and economic success*, 2005, pp. 23–79.

Duncan, Greg J., Chantelle J. Dowsett, Amy Claessens, Katherine Magnuson, Aletha C. Huston, Pamela Klebanov, Linda S. Pagani, Leon Feinstein, Mimi Engel, Jeanne Brooks-Gunn, Holly Sexton, and Kathryn Duckworth, "School readiness and later achievement," *Developmental Psychology*, 2007, *43(6)*, 1428–1446.

Dynarski, Susan, Joshua Hyman, and Diane Whitmore Schanzenbach, "Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion," *Journal of Policy Analysis and Management*, 2013, *32(4)*, 692–717.

Fryer, Roland G. and Steven D. Levitt, "Understanding the black-white test score gap in the first two years of school," *Review of Economics and Statistics*, 2004, *86(2)*, 447–464.

—— **and** —— , "The black-white test score gap through third grade," *American Law and Economics Review*, 2006, *8(2)*, 249–281.

Garces, Eliana, Duncan Thomas, and Janet Currie, "Longer-term effects of Head Start," *American Economic Review*, 2002, *92(4)*, 999–1012.

Gibbs, Chloe, Jens Ludwig, and Douglas L. Miller, "Does Head Start do any lasting good?," Working Paper 17542, National Bureau of Economic Research 2011.

Goodman-Bacon, Andrew, "The long-run effects of childhood insurance coverage: Medicaid implementation, adult health, and labor market outcomes," Working Paper 22899, National Bureau of Economic Research 2016.

Heckman, James J., Rodrigo Pinto, and Peter Savelyev, "Understanding the mechanisms through which an influential early childhood program boosted adult outcomes," *American Economic Review*, 2013, *103(6)*, 2052–2086.

Hertz, Tom, Tamara Jayasundera, Patrizio Piraino, Sibel Selcuk, Nicole Smith, and Alina Verashchagina, "The inheritance of educational inequality: International comparisons and fifty-year trends," *The BE Journal of Economic Analysis & Policy*, 2008, *7* (2), 10.

Hoynes, Hilary W., Diane Whitmore Schanzenbach, and Douglas Almond, "Long-run impacts of childhood access to the safety net," *American Economic Review*, 2016, *106(4)*, 903–934.

Hoynes, Hillary, Marianne Page, and Ann Huff Stevens, "Can targeted transfers improve birth outcomes? Evidence from the introduction of the WIC program," *Journal of Public Economics*, 2011, *95(7-8)*, 813–827.

Johnson, Rucker C. and C. Kirabo Jackson, "Reducing inequality through dynamic complementarity: Evidence from Head Start and public school pending," Working Paper 23489, National Bureau of Economic Research 2017.

Kane, Thomas J, Cecilia Elena Rouse, and Douglas Staiger, "Estimating returns to schooling when schooling is misreported," Technical Report, National bureau of economic research 1999.

Kline, Patrick and Christopher R. Walters, "Evaluating public programs with close substitutes: The case of Head Start," *Quarterly Journal of Economics*, 2016, *131(4).*

Lee, Chul-In and Gary Solon, "Trends in intergenerational income mobility," *Review of Economics and Statistics*, 2009, *91(4)*, 766–772.

Ludwig, Jens and Douglas L. Miller, "Does Head Start improve children's life chances? Evidence from a regression discontinuity design," *Quarterly Journal of Economics*, 2007, *122(1)*, 159–208.

Maurin, Eric and Sandra McNally, "Vive la revolution! Long-term education returns of 1968 to the angry students," *Journal of Labor Economics*, 2008, *26(1)*, 1–33.

Mazumder, Bhashkar, "Fortunate sons: New estimates of intergenerational mobility in the United States using social security earnings data," *Review of Economics and Statistics*, 2005, *87* (2), 235–255.

Montialoux, Claire, "Revisiting the impact of Head Start," Policy Brief, University of California, Berkeley, Institute for Research on Labor and Employment, Berkeley, CA 2016.

Murnane, Richard J., John B. Willett, Kristen L. Bub, and Kathleen McCartney, "Understanding trends in the black-white achievement gaps during the first years of school," in "Brookings-Wharton Papers on Urban Affairs" Brookings Institution Press Washington, DC 2006, pp. 97–135.

Office of Child Development, "Project Head Start 1965-1967: A descriptive report of programs and participants," Report, Department of Health, Education, and Welfare, Washington, DC 1968.

——, "Project Head Start 1968: A descriptive report of programs and participants," Report, Department of Health, Education, and Welfare, Washington, DC 1970.

Oreopoulos, Philip, Marianne E. Page, and Ann Huff Stevens, "The intergenerational effects of compulsory schooling," *Journal of Labor Economics*, 2006, *24(4)*, 729–760.

Page, Marianne, "Fathers' education and children's human capital: Evidence from the World War II G.I. Bill," 2009. mimeo.

Page, Marianne E, "New evidence on the inter generational correlation in welfare participation," *Generational income mobility in North America and Europe*, 2004, p. 226.

Puma, Michael, Stephen Bell, Ronna Cook, Camilla Heid, and et al., "Head Start impact study: Final report," Report, U.S. Department of Health and Human Services, Administration for Children and Families, Washington, DC 2010.

___, ___, ___, ___, Michael Lopez, and et al., "Head Start impact study: First year findings," Report, U.S. Department of Health and Human Services, Administration for Children and Families, Washington, DC 2005.

___, ___, ___, ___, Pam Broene, Frank Jenkins, Andrew Mashburn, and Jason Downer, "Third grade follow-up to the Head Start impact study: Final report," Report 2012-45, U.S. Department of Health and Human Services, Administration for Children and Families, Washington, DC 2012.

Reardon, Sean F. and Ximena A. Portilla, "Recent trends in income, racial, and ethnic school readiness gaps at kindergarten entry," *AERA Open*, 2016, *2(3)*, 1–18.

Rossin-Slater, Maya and Miriam Wust, "What is the added value of preschool? Long-term impacts and interactions with a health intervention," Working Paper 22700, National Bureau of Economic Research 2016.

Sawhill, Isabel V., Scott Winship, and Kerry Searle Grannis, "Pathways to the middle class: Balancing personal and public responsibilities," Report, Brookings Institution, Washington, DC 2012.

Schweinhart, Lawrence J., Jeanne Montie, Zongping Xiang, W. Steven Barnett, Clive R. Belfield, and Milagros Nores, *Lifetime Effects: The High/Scope Perry Preschool Study Through Age 40*, Ypsilanti, MI: High/Scope Press, 2005.

Solon, Gary, "Intergenerational mobility in the labor market," *Handbook of labor economics*, 1999, *3*, 1761–1800.

Thompson, Owen, "Head Start's impact in the very long-run," 2017. Unpublished paper.

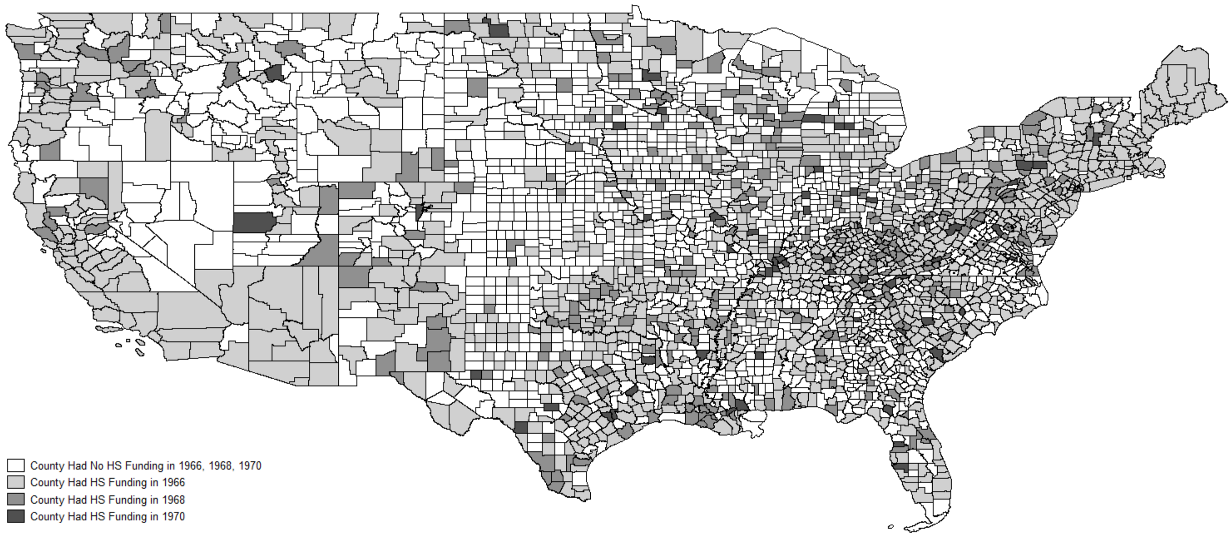
Vinovskis, Maris A., *The Birth of Head Start*, Chicago, IL: University of Chicago Press, 2005.

Walters, Christopher R., "Inputs in the production of early childhood human capital: Evidence from Head Start," *American Economic Journal: Applied Economics*, 2015, *7(4)*, 76–102.

Zigler, Edward and Jeanette Valentine, *Project Head Start: A Legacy of the War on Poverty*, New York, NY: Free Press, 1979.

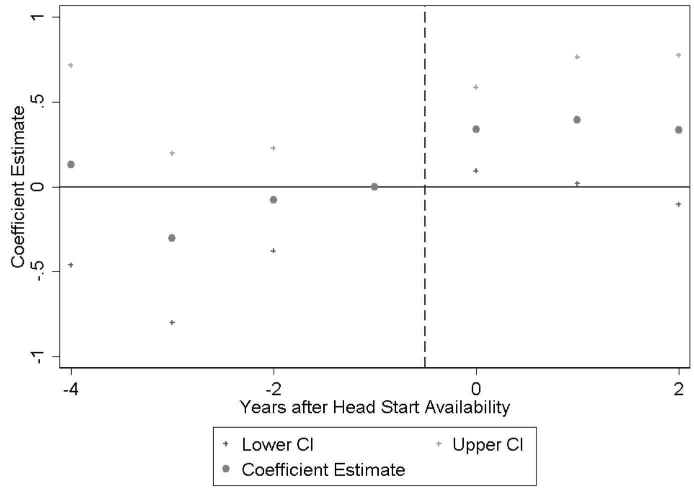
Figures

Figure 1: Early Geographic Expansions of Head Start



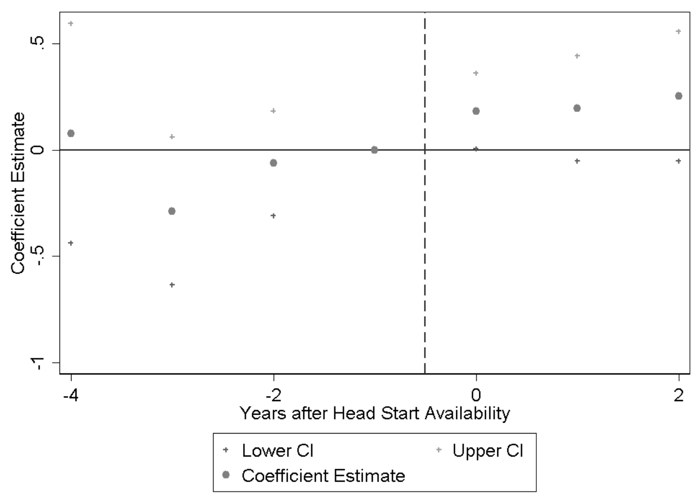
Source: National Archives and Records Administration.

Figure 2: Event Study for Second-Generation Index (High Impact)



Note: Circles indicate coefficients on indicator variables for the difference between a mothers birth cohort and the first birth cohort to have Head Start availability within a county (non-negative values reflect cohorts with Head Start availability). The dependent variable is the index of adult outcomes in the second generation. Regressions include cohort and county of birth fixed effects. Standard errors are clustered at the birth county level. Asterisks indicate 95 percent confidence intervals.

Figure 3: Event Study for Second-Generation Index (Low Impact)



Note: Circles indicate coefficients on indicator variables for the difference between a mothers birth cohort and the first birth cohort to have Head Start availability within a county (non-negative values reflect cohorts with Head Start availability). The dependent variable is the index of adult outcomes in the second generation. Regressions include cohort and county of birth fixed effects. Standard errors are clustered at the birth county level. Asterisks indicate 95 percent confidence intervals.

Tables

Table 1: Descriptive Statistics of NLSY79 Women and Their Children

Sample	(1) (Full Sample)	(2) High Impact (Grandmother <HS)	(3) Low Impact (Grandmother ≤HS)
Second Generation (Child) Outcomes			
Teen Parent	0.19	0.22	0.19
Crime	0.29	0.31	0.30
High School	0.82	0.78	0.82
Some College	0.58	0.51	0.56
Index	0.00	-0.12	-0.01
Observations	3533	1687	2732
First Generation (Mother) Outcomes			
Teen Mom	0.32	0.52	0.43
High School	0.86	0.83	0.86
Some College	0.47	0.33	0.41
Highest Grade	13.2	12.5	12.8
Income	46919	34534	40123
Observations	2398	821	1398

Note: The bottom panel presents sample means for women in the NLSY79. The top panel provides sample means for the children of these women, restricted to individuals over 20 in 2012. Each column provides sample means for a different sample, corresponding to a particular research design and set of sample restrictions. Column (1) provides sample means for NLSY79 women and their children. Columns (2) and (3) provide analogous means for the two samples used with our preferred research design, the changing availability of Head Start within counties. Each of these samples is restricted based on the education level of the mother of the NLSY79 participant (i.e., the grandmother of the children). Column (2), our "high impact" sample is restricted to participants in the NLSY79 whose mothers dropped out of high school. Column (3), our "low impact" sample, is restricted to participants in the NLSY79 whose mothers finished high school but completed no additional education. We largely follow Deming (2009) in our construction of crime and income measures. Crime is defined as any arrests, convictions, or probations. Income is a measure of the mother's permanent lifetime income, calculated as the deflated average of net family income for the mother. To address multiple inference concerns and reduce measurement error, we follow the prior literature in constructing a summary index of our outcome measures (Kling, Liebman, and Katz 2007; Deming 2009). We normalize each outcome to have a mean zero and standard deviation one, adjust outcome signs so that a more positive outcome is better (i.e., we flip the sign on teen parenthood and crime), and take the simple average across these outcomes. We then standardize this summary index.

Table 2: Geographic Variation: Reduced Form Effect of Head Start in County

VARIABLES	(1) Teen Parent	(2) Crime	(3) High School	(4) Some College	(5) Index
Grandmother $<$ High School	-0.086***	-0.156***	0.127***	0.169***	0.466***
	(0.031)	(0.043)	(0.048)	(0.055)	(0.101)
Mean	0.22	0.30	0.77	0.52	-0.12
Observations	1,709	1,709	1,709	1,687	1,687
Grandmother $\leq$ High School	-0.058**	-0.063*	0.064	0.070	0.218**
	(0.028)	(0.034)	(0.039)	(0.045)	(0.085)
Mean	0.19	0.30	0.81	0.56	-0.01
Observations	2,770	2,770	2,769	2,732	2,732

Note: Each column represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. The dependent variables are second generation outcomes as indicated by the column titles. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include child gender, age, and age squared, as well as mother's birth order and race. * ($p<0.10$) **($p<0.05$), ***($p<0.01$).

Table 3: Geographic Variation: Reduced Form Effect of Head Start in County (robustness)

VARIABLES	(1) Index	(2) Index	(3) Index	(4) Index	(5) Index	(6) Index	(7) Index
Grandmother < High School	0.417***	0.466***	0.436***	0.488***	0.370**	0.292*	0.384*
	(0.111)	(0.101)	(0.105)	(0.105)	(0.144)	(0.167)	(0.230)
Observations	1,688	1,687	1,687	1,653	1,687	1,687	1,687
Grandmother $\leq$ High School	0.196**	0.218**	0.221***	0.234***	0.213**	0.243**	0.093
	(0.087)	(0.085)	(0.082)	(0.086)	(0.108)	(0.109)	(0.166)
Observations	2,733	2,732	2,732	2,669	2,732	2,732	2,732
Covariates	N	Y	Y	Y	Y	Y	Y
Birth County Chars. (1960) x Trend	N	N	Y	N	N	N	N
WOP Measures	N	N	N	Y	N	N	N
Birth County Trends	N	N	N	N	Y	N	N
State by Year Fixed Effects	N	N	N	N	N	Y	N
Family Fixed Effects (mother)	N	N	N	N	N	N	Y

Note: Each column represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. The dependent variable is the index of second generation outcomes as indicated by the column titles. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include child's gender, age, and age squared, as well as mother's birth order and race. Column (2) contains base specification. * (p<0.10) **(p<0.05), ***(p<0.01).

Table 4: Geographic Variation: Reduced Form Effect of Head Start in County (falsification)

VARIABLES	(1) Teen Parent	(2) Crime	(3) High School	(4) Some College	(5) Index
Grandmother $\geq$ High School	-0.028	0.020	-0.015	0.001	-0.021
	(0.041)	(0.053)	(0.045)	(0.069)	(0.110)
Mean	0.14	0.27	0.88	0.68	0.22
Observations	1,355	1,355	1,354	1,338	1,338

Note: Each column represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. Rows indicate subsample based on selection of mother's birth years. The dependent variables are second generation outcomes as indicated by the column titles. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include child's gender, age, and age squared, as well as mother's birth order and race. Sample restricted to mothers whose mothers had at least a high school degree. * (p<0.10) **(p<0.05), ***(p<0.01).

Table 5: Geographic Variation: Reduced Form Effect of Head Start in County (Mother)

VARIABLES	(1) Teen Parent	(2) High School	(3) Some College	(4) Highest Grade	(5) Perm Income (std)	(6) Number of Kids
Grandmother < High School	0.030	0.110**	0.092*	0.571***	0.162*	-0.136
	(0.072)	(0.048)	(0.053)	(0.218)	(0.089)	(0.133)
Mean	0.52	0.81	0.32	12.4	-0.217	2.74
Observations	824	821	821	821	824	824
Grandmother ≤ High School	0.043	0.059*	0.060	0.470**	0.050	0.082
	(0.048)	(0.036)	(0.043)	(0.184)	(0.068)	(0.108)
Mean	0.43	0.85	0.40	12.8	-0.075	2.62
Observations	1,407	1,398	1,398	1,398	1,407	1,407

Note: Each column represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. The dependent variables are first-generation (mother) outcomes indicated by the column titles. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include mother's birth order and race. * ($p<0.10$) **($p<0.05$), ***($p<0.01$).

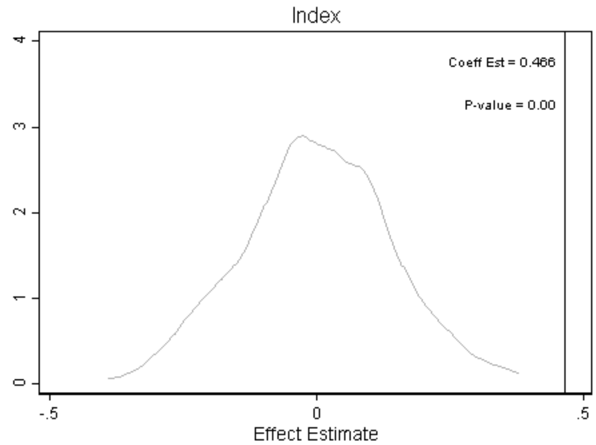
Table 6: Geographic Variation: Reduced Form Effect of Head Start in County (heterogeneity)

VARIABLES	(1) Teen Parent	(2) Crime	(3) High School	(4) Some College	(5) Index
Black & Grandmother $<$ High School	-0.114**	-0.086**	0.125	0.113	0.396**
	(0.053)	(0.041)	(0.079)	(0.080)	(0.153)
Black & Grandmother $\leq$ High School	-0.058	-0.011	0.089	0.050	0.185
	(0.048)	(0.042)	(0.063)	(0.078)	(0.137)
Male & Grandmother $<$ High School	-0.015	-0.271***	0.144*	0.147**	0.505***
	(0.045)	(0.066)	(0.078)	(0.073)	(0.129)
Male & Grandmother $\leq$ High School	-0.043	-0.139***	0.102*	0.082	0.321***
	(0.040)	(0.052)	(0.060)	(0.058)	(0.106)
South & Grandmother $<$ High School	-0.103**	-0.101**	0.155**	0.186**	0.491***
	(0.048)	(0.043)	(0.068)	(0.075)	(0.130)
South & Grandmother $\leq$ High School	-0.076*	-0.054	0.097*	0.092	0.286**
	(0.040)	(0.037)	(0.055)	(0.072)	(0.116)

Note: Each column represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's county of birth. Rows indicate subsample (South designation is based on mother's birth county). The dependent variables are second generation outcomes as indicated by the column titles. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include child's gender, age, and age squared, as well as mother's birth order and race. * (p<0.10) **(p<0.05), ***(p<0.01).

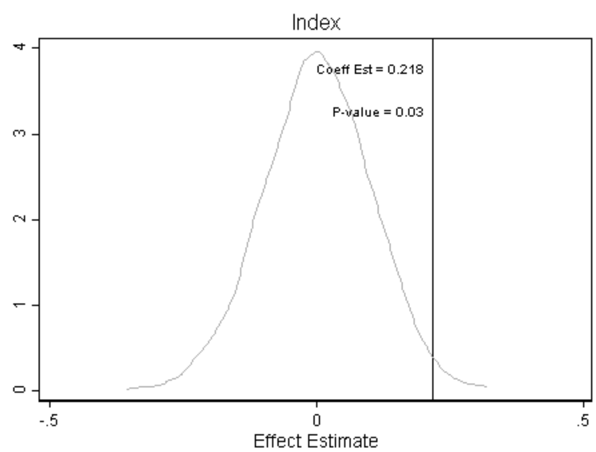
Appendix A: Supplemental Figures and Tables

Figure A1: High Impact Sample: Randomization Inference (index)



Note: The figure plots the smoothed distribution of coefficient estimates of Head Start availability for 1000 random assignments of the pattern of timing of the Head Start rollout to counties. The vertical line indicates the coefficient estimate using the actual timing of Head Start availability in each county. P-value presented is the two-tailed statistic calculated as the share of coefficient estimates obtained under random assignment of Head Start introduction timing that are larger in absolute magnitude than the estimate using the actual timing of introduction.

Figure A2: Low Impact Sample: Randomization Inference (index)



Note: The figure plots the smoothed distribution of coefficient estimates of Head Start availability for 1000 random assignments of the pattern of timing of the Head Start rollout to counties. The vertical line indicates the coefficient estimate using the actual timing of Head Start availability in each county. P-value presented is the two-tailed statistic calculated as the share of coefficient estimates obtained under random assignment of Head Start introduction timing that are larger in absolute magnitude than the estimate using the actual timing of introduction.

Table A1: Geographic Variation: Head Start Availability and Enrollment

	Grandmother $<$ HS	Grandmother $\leq$ HS
1994 Retrospective Self-reported Head Start Participation (NLSY 79)		
Head Start in County	0.100**	0.055
	(0.045)	(0.034)
Observations	805	1,374
Mean Participation | Head Start not in County	0.14	0.12
Mean Participation | Head Start in County	0.35	0.30
Fraction Enrolled in Head Start in State (OEO 66 Enrollment Counts)		
Head Start in County	0.287*	0.149*
	(0.150)	(0.077)
Observations	49	49
Implied Mean Fraction Enrolled | Head Start in County	0.58	0.30
Fraction Enrolled in Head Start Nationally (Enrollment Counts (66-69)		
Observations	4	4
Implied Mean Fraction Enrolled | Head Start in County	0.56	0.29

Note: Each cell represents a separate OLS regression. The first panel contains estimates of the effect of Head Start availability on self-reported participation in the NLSY79 (reported in 1994). The dependent variable is the mother's self-reported Head Start status. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include mother's birth order and race. Sample restricted to mothers whose mothers had less than a high school degree or at most a high school degree. The means are average self-reported Head Start participation levels in counties with and without Head Start availability. The second panel contains estimates from a regression of the state-level share of four-year olds participating in Head Start against the fraction of the state's four-year old population with Head Start availability. State level availability is determined by dividing the number of 4 year-olds born in counties with availability in 1966 (determined using our county-level treatment variable and natality data) by the total number of 4 year-olds born in the state. State-level participation and availability counts are adjusted using statistics on the education levels of the mothers of participants (70% had less than a high school degree and 90% had a high school degree or less) and the mothers of children born during this time period (33% had less than a high school degree and 82% had a high school degree or less). The means are the fraction of children estimated to be enrolled in Head Start, assuming that only children born in counties with Head Start availability are enrolled. The third panel contains similar means using the national level Head Start enrollment data for 1966-1969 combined with natality data. * (p<0.10) **(p<0.05), ***(p<0.01).

Table A2: Geographic Variation Balance Checks: Head Start Availability and Observables

VARIABLES	(1) First born	(2) Black	(3) HH in Poverty (78)	(4) Grandmother $\leq$ High School	(5) Grandmother $<$ High School	(6) Child Age (12)	(7) Male Child	(8) Predicted Index
Full Sample	0.060	0.045	0.020	0.026	-0.001	0.036	-0.001	-0.028
	(0.042)	(0.029)	(0.047)	(0.031)	(0.039)	(0.250)	(0.035)	(0.026)
Mean	0.50	0.35	0.33	0.90	0.54	27.15	0.50	0.01
Observations	1,676	1,685	1,595	1,585	1,585	3,272	3,273	2,917
Grandmother $<$ HS	0.045	0.021	-0.015	N/A	N/A	0.096	0.116***	-0.033
	(0.063)	(0.044)	(0.077)			(0.338)	(0.043)	(0.035)
Mean	0.48	0.38	0.45	1	1	27.72	0.50	-0.12
Observations	824	826	792	826	826	1,709	1,710	1,641
Grandmother $\leq$ HS	0.067	0.045	0.022	N/A	-0.011	0.229	0.026	-0.028
	(0.048)	(0.032)	(0.052)		(0.042)	(0.261)	(0.035)	(0.025)
Mean	0.50	0.35	0.35	1	0.60	27.27	0.50	-0.03
Observations	1,407	1,412	1,347	1,412	1,412	2,770	2,771	2,650

Note: Each cell represents a separate OLS regression with robust standard errors in parentheses, clustered on county of birth. The dependent variable is indicated by the column title. The generate the predicted index we first regress our child index of adult outcomes on all of the maternal and child characteristics in the table. Coefficients presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. Regressions also include controls for year of birth and county of birth fixed effects. * ($p<0.10$) **($p<0.05$), ***($p<0.01$).

Table A3: Geographic Variation: Head Start Availability and Self-reported Participation (Falsification)

	Grandmother $\geq$ High School
HS in County	-0.005
	(0.053)
Observations	735
Mean	0.15

Note: Each column represents a separate OLS regression with robust standard errors in parentheses, clustered on county of birth. The dependent variable is the mother's self-reported Head Start status. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include mother's birth order and race. Sample restricted to mothers whose mothers had at least a high school degree. * (p<0.10) **(p<0.05), ***(p<0.01).

Appendix B: Measurement Error and the Relationship Between Head Start Availability and Retrospective Self-Reported Participation

Table A1 illustrates that our measure of Head Start availability predicts both self-reported Head Start participation and state-level participation rates derived from administrative Head Start enrollment data. The top panel presents estimates of the effect of Head Start availability on self-reported participation. When a program is available in a county four or five years after a mother's year of birth, the mother in our high impact sample is 10 percentage points more likely *to report* having participated in Head Start as a child. Given the self-reported and retrospective nature of the Head Start participation variable, we expect there is considerable misreporting. As has been established in the literature, measurement error in a binary variable is necessarily non-classical and will thus result in a downward biased estimate of the relationship between Head Start availability and participation. For the purposes of explanation, consider the following specification of the true relationship between Head Start participation HS and Head Start availability $HSavail$:

$$HS = \beta_1 X + \beta_2 HSavail + \varepsilon \tag{2}$$

Here, HS is an individuals true Head Start participation, X is a vector of other observables (in practice this includes year and county of birth fixed effects), and $HSavail$ is an indicator for the availability of Head Start. Instead of observing HS, we observe self-reported Head Start participation HS_{sr}, which is HS measured with error. Given that HS is a binary variable, misclassification will necessarily lead to non-classical measurement error and a downward biased estimate of β_2 (Kane et al. 1999, Aigner 1973).

The extent of this bias has been previously derived. Let $m_1 = Pr(HS_{sr} = 0|HS = 1)$ represent the fraction of individuals who self-report not participating in Head Start when they did participate. Similarly, let $m_0 = Pr(HS_{sr} = 1|HS = 0)$ represent the fraction of individuals who self-report participating in Head Start when they did not participate. Under the assumption of constant rates of misclassification (m_1 and m_0), $\beta_1^{OLS} = (1m_0m_1) * \beta_1$.

Determining the extent of the bias depends on knowledge of the misclassification rates m_1 and m_0. Because we do not have measures of these misclassification rates, we take two approaches to guide the selection of reasonable choices for m_1 and m_0. First, we use the available administrative and self-reported participation data to generate estimates of the misclassification rates. Second, we draw upon the existing literature on retrospective recall rates to inform our selection of misclassification rates.

Data-driven Estimates of Misclassification Rates

For the data driven exercise, we draw upon the following equations:

$$\begin{aligned}(\overline{HS}_{sr}|HSavail = 1) &= Pr(HS_{sr} = 1|HS = 0) * Pr(HS = 0|HSavail = 1) \\ &+ Pr(HS_{sr} = 1|HS = 1) * Pr(HS = 1|HSavail = 1)\end{aligned}$$

and,

$$\begin{aligned}(\overline{HS}_{sr}|HSavail = 0) &= Pr(HS_{sr} = 1|HS = 0) * Pr(HS = 0|HSavail = 0) \\ &+ Pr(HS_{sr} = 1|HS = 1) * Pr(HS = 1|HSavail = 0)\end{aligned}$$

From the data, we can calculate the left hand side of both equations. If we assume that Head Start participation occurs entirely in counties with Head Start funding, we can set $Pr(HS = 1|HSavail = 0)$ to 0 and $Pr(HS = 0|HSavail = 0)$ to 1, and use the administrative data on Head Start participation levels to provide estimates of $Pr(HS = 0|HSavail = 1)$ and $Pr(HS = 1|HSavail = 1)$. This simplifies both equations to:

$$\begin{aligned}(\overline{HS}_{sr}|HSavail = 1) &= Pr(HS_{sr} = 1|HS = 0) * Pr(HS = 0|HSavail = 1) \\ &+ Pr(HS_{sr} = 1|HS = 1) * Pr(HS = 1|HSavail = 1)\end{aligned}$$

and,

$$(\overline{HS}_{sr}|HSavail = 0) = Pr(HS_{sr} = 1|HS = 0) = m_0$$

We now have two equations and two unknowns. From the data on our high impact sample, we know that $(\overline{HS}_{sr}|HSavail = 0) = m_0$ is somewhere between 0.20 and 0.24. Plugging in our population-level averages for Head Start participation in counties with Head Start availability (0.58) and solving provides bounds for m_1 of 0.54 to 0.57. Combining this with the equation for $(\overline{HS}_{sr}|HSavail = 1)$, indicates that the relationship between Head Start availability and participation is likely biased down by a factor of 3.9 to 5.3. In other words, under a reasonable set of assumptions, the true magnitude of β_1 is closer to 39 to 53 percentage points. The true estimate may be even higher if Head Start availability increases rates of misclassification.

Literature-driven Selection of Misclassification Rates

A second way to guide the selection of reasonable choices for m_1 and m_0 is to draw upon existing literature on retrospective recall rates. While numerous studies in the psychology literature indicate that recall of early life events is often quite poor, there are only a handful of studies that have managed to track individuals over time so as to compare actual or reported events during childhood or adolescence with retrospective information regarding the same events. These studies all indicate the poor reliability of adulthood retrospective accounts of childhood events across a series of outcomes including various forms of childhood abuse, parental divorce or separation, parental chronic unemployment, household legal trouble, household illness or disability, maternal depression, household substance abuse, household incarceration, residence changes, height, weight, injuries, and contacts with the criminal justice system (Henry et al. 1994, Naicker et al. 2017; Reuben et al. 2016). Even parental death, which is recalled far more accurately than these other measures, is imperfectly measured in retrospective reports.

While several studies indicate that retrospective accounts are fraught with misclassification issues, only one contains the necessary underlying data to easily construct measures of $\frac{1}{1-m_0-m_1}$ (Naicker et al. 2017). Naicker and coauthors (2017) compare the reports of adolescents at ages 11, 15, and 18 with retrospective reporting by the same individuals at age 23 across a variety of events. The mean and median measures $m_0 + m_1$ are 0.77 to 0.88, suggesting a scaling factor between 4.35

and 8.36. Using these measures would imply a first-stage relationship of 43.5 to 83.6 percentage points. It is also important to note that the gap here between event and recall is substantially shorter than the 30 to 35 year gap between the year of likely Head Start participation and the time the retrospective question was asked (1994) in the NLSY79. This suggests that recall may be even worse, and thus the scaling factor should be even higher, in the NLSY. Reuben et al. (2016) ask retrospective questions to a much older group of individuals, age 38, but focuses on recall of a similar set of childhood experiences to those explored by Naicker and colleagues (2017). This gap between the timing of the event and the timing of recall is much closer to that observed in the NLSY79. While the data provided do not allow for the construction of misclassification rates, Cohens Kappa values, which are constructed in both studies, are similarly very low, indicating poor agreement between retrospective and contemporaneous reporting.

Endnotes

Henry, Bill, Terrie E. Moffitt, Avshalom Caspi, John Langley, and Phil A. Silva, "On the "remembrance of things past": A longitudinal evaluation of the retrospective method," *Psychological Assessment*, 2016, *6*, 92-101.

Naicker, Sara N., Shane A. Norris, Musawenkosi Mabaso, and Linda M. Richter, "An analysis of retrospective and repeat prospective reports of adverse childhood experiences from the South African Birth to Twenty Plus cohort" *PloS ONE*, 2017, *12(7)*, 1-19.

Reuben, Aaron, Terrie E. Moffitt, Avshalom Caspi, Daniel W. Belsky, and et al., "Lest we forget: comparing retrospective and prospective assessments of adverse childhood experiences in the prediction of adult health," *Journal of Child Psychology and Psychiatry*, 2016, *57(10)*, 1103-1112.

Table B1: Geographic Variation: Head Start Availability and Enrollment

	Grandmother < HS	Grandmother $\leq$ HS
1994 Retrospective Self-reported Head Start Participation (NLSY 79)		
Head Start in County	0.100**	0.055
	(0.045)	(0.034)
Observations	805	1,374
Mean Participation | Head Start not in County	0.14	0.12
Mean Participation | Head Start in County	0.35	0.30
Fraction Enrolled in Head Start in State (OEO 66 Enrollment Counts)		
Head Start in County	0.287*	0.149*
	(0.150)	(0.077)
Observations	49	49
Implied Mean Fraction Enrolled | Head Start in County	0.58	0.30
Fraction Enrolled in Head Start Nationally (Enrollment Counts (66-69)		
Observations	4	4
Implied Mean Fraction Enrolled | Head Start in County	0.56	0.29

Note: Each cell represents a separate OLS regression. The first panel contains estimates of the effect of Head Start availability on self-reported participation in the NLSY79 (reported in 1994). The dependent variable is the mother's self-reported Head Start status. Coefficient presented for binary variable indicating Head Start availability in mother's birth county 4 or 5 years after the year of mother's birth. In addition to year of birth and county of birth fixed effects, controls include mother's birth order and race. Sample restricted to mothers whose mothers had less than a high school degree or at most a high school degree. The means are average self-reported Head Start participation levels in counties with and without Head Start availability. The second panel contains estimates from a regression of the state-level share of four-year olds participating in Head Start against the fraction of the state's four-year old population with Head Start availability. State level availability is determined by dividing the number of 4 year-olds born in counties with availability in 1966 (determined using our county-level treatment variable and natality data) by the total number of 4 year-olds born in the state. State-level participation and availability counts are adjusted using statistics on the education levels of the mothers of participants (70% had less than a high school degree and 90% had a high school degree or less) and the mothers of children born during this time period (33% had less than a high school degree and 82% had a high school degree or less). The means are the fraction of children estimated to be enrolled in Head Start, assuming that only children born in counties with Head Start availability are enrolled. The third panel contains similar means using the national level Head Start enrollment data for 1966-1969 combined with natality data. * ($p<0.10$) **($p<0.05$), ***($p<0.01$).

Appendix C: Data Appendix

To generate measures of Head Start exposure in the late 1960s, we compile data from the National Archives and Records Administration (NARA) files on the Office of Economic Opportunity's Head Start grant funding (National Archives, n.d.). We employ two data sources, covering different spans of time, to construct county-level measures of Head Start activity.

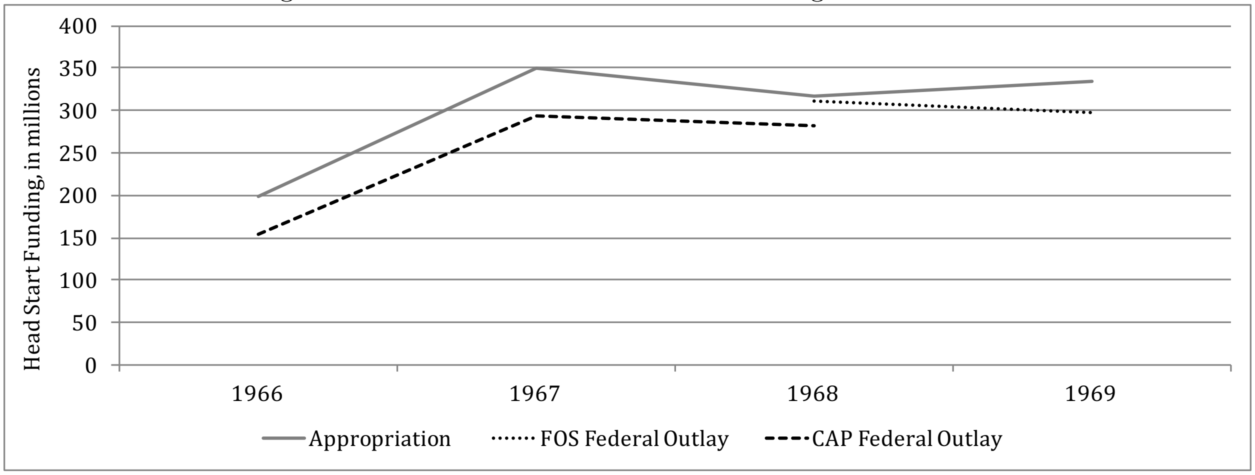
Community Action Program (CAP) Records, 1966–1968

The first data source is Records About Community Action Program (CAP) Grants and Grantees, NARA Record Group 381 (National Archives identifier 604417). These CAP files, created by the Community Services Administration, document the timeframe July 1964 through September 1981. The CAP records consist of two files, the Grantee Organization Master Files which provide information about the grantees receiving funds through CAP and the Funded Program Account Master Files which document specific grant actions. For the purposes of creating county-level funding amounts, we use the Funded Program Account Master Files which contain state and county identifiers.

We retain only records for which the grant action is Initial Application, Supplemental Funding, or Refunding at End of Program Year or for which the grant action field is blank. We exclude, therefore, grant actions Deobligation, Extension, Transfer In, and Transfer Out.

These files contain grant actions on a variety of CAP-funded poverty programs, including job training, housing services, health services, and community development. For this reason, we extract files with certain search terms in the project description field: Head Start, Headstart, child dev, preschool, pre-school, early childhood, HS child, and FYHS. We do not use the terms child care, daycare, or family care center though they appear in the project description field. These filters result in retaining only a subset of grant actions for each grantee number. Once we have the domain of Head Start programs, we aggregate federal funding to the county-by-FY level using state and county geographic codes. Notably, we drop nonnumeric characters from the funding amount field when they appear (always at the end of the field), assuming that these are placeholders for an input mask.

Figure C1: Crosswalk of Head Start Funding Data Sources



Source: National Archives Records Administration, Records of the Community Services Administration & U.S. Department of Health and Human Services, Administration for Children and Families, Office of Head Start.

Because Head Start data is missing from the electronic files for FY 1965 (also documented in Bailey and Duquette's (2014) data appendix), we construct CAP records for FY 1966, 1967, and 1968. After 1968, the search terms we employ largely drop out of the project description field and no longer appear after 1972. While we undercount total Head Start grant funding relative to published federal appropriations (Office of Head Start 2015), the pattern of funding levels and changes across these three years tracks well, as displayed in Figure C1.

Federal Outlays System (FOS) Files, 1968–1980

The second source is the Federal Outlays System (FOS) Files, also NARA Record Group 381 (National Archives identifier 599052). These records were collected from July 1967 through September 1980, also by the Community Services Administration. The files contain data on Federal Executive Branch outlays and include four files for each fiscal year: 1) a County/State Master File, 2) a City Master File, 3) a Geographic Table File, and 4) a Program Appropriations, Functions, and Agency Table File. The County/State Master Files for each year contain information on programs and outlays with state and county identifiers. We compile these records with the Program

Appropriations, Functions, and Agency Table File across the years and again employ search terms to narrow to Head Start programs.

These files contain a variety of program types, so we search in the program title field for the following terms: Head Start, Headstart, child dev, early childhood, HS child, FYHS, and follow-thru program (OEO). Some terms are employed to align with Ludwig and Miller's (2007) file creation process. Terms related to preschool were excluded as they captured unrelated school-age programs. Records containing the following terms in the title field (appearing in conjunction with child dev or early childhood) were dropped: handicapped, handic, child abuse prev, and child welfare. In addition, 78 observations were deleted because they were missing state identifiers. Funding outlays were then aggregated to the county-by-FY level.

Appendix D: Family Fixed Effects Approach

Family Fixed Effects Approach

Given discussion of the likelihood of extensive mismeasurement of the self-reported Head Start participation measure, we place little credibility in the family fixed effects approach. However, for completeness we show results from a family fixed effects model where we compare the outcomes of children of sisters who differ in their Head Start participation. This model is very similar to that used previously to explore effects of Head Start on the participants (Deming 2009, Garces et al. 2002, Currie and Thomas 1995). The basic specification is:

$$y_{ij} = \alpha + \beta X_i + \theta HeadStart_i + \gamma_j + \varepsilon_{ij} \tag{3}$$

where y_{ij} is an adult outcome for a child; X_i includes controls for the child's sex and age as well as the mother's birth year and birth order; and γ_j are mother's family fixed effects. $HeadStart_i$ captures participation in Head Start as a binary variable, turning on for the affected sister within families. Robust standard errors are clustered on the mother's 1979 household. We are interested in identifying θ, the effect of Head Start participation on the adult outcomes of the participants' children. Given prior discussion regarding the likely magnitude of misclassification in Head Start participation, we consider estimates of θ to be merely suggestive.

Our baseline results in Table D1 indicate a significant (0.26 sd) improvement in the summary index of adult outcomes of a mother's children when the mother reports having participated in Head Start as a child. While the estimates for the component outcomes are not statistically different from zero, the magnitude and direction of all of the coefficients suggest improvements in outcomes, with a 3 percentage point reduction in the share of children who become parents as teenagers, a 9 percentage point reduction in criminal behavior, and 7 percentage point increases in the share of individuals who graduate from high school and attend college. Following the discussion in Appendix B, the likely magnitude of the misclassification in Head Start participation indicates that these estimates are substantially downwardly biased.

Furthermore, as with all studies that leverage family fixed effect designs, there are questions about what generates variation within a family in program participation (or reported participation). For example, it is possible that the parents of NLSY79 participants placed greater focus on the education or care of one child versus another, perhaps as a result of changing financial security or perceived differences in aptitude. This would result in overestimates of the effect of Head Start. On the other hand, Head Start eligibility depended, and continues to depend, largely on family resources; at least 90 percent of Head Start participants served at each program site had to be from families below the federal poverty line. This constraint resulted in within family variation in eligibility that was correlated with family resources. In other words, when comparing a sibling who participated with one who did not, the one who participated is more likely to have grown up in poverty than the one who did not participate. We would expect this to bias our estimates of the effect of Head Start participation down.[29]

Given concerns related to the measurement and endogeneity of participation we place little weight on these results and report them primarily for completeness. The estimation strategy employed in the paper circumvents both of these concerns by using plausibly exogenous variation in Head Start availability.

[29]Spillovers between siblings would also lead to underestimates of the effect of participation.

Table D1: Family FEs: Effect of Head Start Participation on Next Generation Outcomes

VARIABLES	(1) Teen Parent	(2) Crime	(3) High School	(4) Some College	(5) Index
Mother Head Start	-0.026	-0.101	0.092	0.073	0.263*
	(0.065)	(0.077)	(0.065)	(0.076)	(0.148)
Mean	0.27	0.33	0.78	0.50	-0.19
Observations	3,580	3,580	3,579	3,533	3,533

Note: Each column represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's 1979 household. The dependent variables are second-generation outcomes as indicated by the column titles. Coefficient presented for binary variable indicating first-generation (mother's) self-reported participation in Head Start. In addition to family fixed effects, controls include child's birth order, sex, and age, and mother's birth order and age. Sample includes individuals over 20 in 2012. There is insufficient variation within the sibling comparisons to include mother or child's race/ethnicity.* (p<0.10) **(p<0.05), ***(p<0.01).

Table D2: Family FEs: Effect of Head Start Participation on Next Generation Outcomes (heterogeneity)

VARIABLES	(1) Teen Parent	(2) Crime	(3) High School	(4) Some College	(5) Index
Black	-0.090	-0.094	0.087	0.055	0.305
	(0.081)	(0.108)	(0.087)	(0.106)	(0.209)
Male	-0.024	-0.220*	0.195*	0.203	0.543***
	(0.066)	(0.121)	(0.102)	(0.130)	(0.205)
South	0.030	-0.213	0.033	0.160	0.348
	(0.113)	(0.157)	(0.106)	(0.146)	(0.287)

Note: Each column represents a separate OLS regression with robust standard errors in parentheses, clustered on mother's 1979 household. Rows indicate subsample (South designation is based on mother's birth county). The dependent variables are indicated by the column titles. Coefficient presented for binary variable indicating mother's self-reported participation in Head Start. In addition to family fixed effects, controls include child's birth order, sex, and age, and mother's birth order and age. There is insufficient variation within the sibling comparisons to include mother or child's race/ethnicity.* ($p<0.10$) **($p<0.05$), ***($p<0.01$).